DOI: 10.13009/EUCASS2019-282

Filter design using an iterative Hilbert transform scheme

*Philippe SAUNOIS**
**ArianeGroup, Les Mureaux, France*
Email: philippe.saunois@ariane.group

Abstract

This paper presents a filtering approach based on an iterative Hilbert transform scheme. The objectives are twice: first to find out what is achievable in terms of filtering satisfying gain and phase gauges with minimum phase lag at low frequencies whatever the order of the filter, secondly to develop insight by showing through the iterations how the Bayard-Bode constraints act on the geometry of the filter transfer function.
Usually, the order of the filter is taken as an input and the filtering problem is solved in a space whose dimension is imposed by the filter order. Here the problem is addressed in two steps. The first step is aiming at finding a solution in a wider space considering only analyticity constraints. Those constraints are completely captured by the iterative Hilbert transform scheme. Then the second step consists in approximating the results by a fixed order filter.
Originally designed in order to initialize a descent optimization scheme with a solution close to optimum, the results of the process have turned out to be sufficient on their own. The approach can be easily tailored to a wide variety of filtering objectives, such as mixed gain and phase constraints either for different weakly damped dynamics or for one dynamic with uncertainties.

1. Introduction

Modern control tools are mainly based on algebra. For sure they have turned out to be very powerful but it can be argued that algebra is far from John A. Wheeler's First Moral Principle [1]: "Never make a calculation without knowing the answer." In order to develop insight, representation is at stake and a geometrical approach is well suited. In that prospective, the paper presents a filtering approach based on an iterative Hilbert transform scheme. It is applied to the design of a linear filter satisfying gain and phase gauges with minimum phase lag at low frequencies. That classical filter design problem is usually solved with a constraint on the filter order. Such a constraint helps to reduce the size of the filter parameters space which is of great help using an algebraic approach. Here the problem is addressed in two steps.

The first step is aiming at finding a solution in a wider space considering analyticity constraints only. Those constraints are completely captured by the iterative Hilbert transform scheme. Then the second step consists in approximating the results by a fixed order filter. By the way, it is possible to answer the classical question when a filter is not fully satisfactory: “what would be the benefit to increase the filter order?” Anyone who has already tuned a filter knows at least from practical experience that the answer is “not much” above some rather low order… Considering analyticity constraints only, it is possible to state that the phase lag introduced by the filter at low frequencies cannot be less than some value whatever the filter order. The approach is detailed in §2.
Three application examples are detailed in §3, focusing on the dynamical evolution of the Bayard-Bode constraints on the Bode diagram flowing from most constrained frequency bands to less constrained frequency bands. This process helps to really “understand” the final result.

The second step is based on a classical RMS approximation with weighting functions: as shown in §4, the good point is that an analytical transfer can be easily approximated with fixed low order filters. This can however introduce some small deviations with respect to the gain and phase gauges and that work was originally performed in order to initialize a descent optimization scheme with a solution close to optimum. That part is not detailed here as the results of the process have turned out to be sufficient on their own.
Some minor limitations of the approach are discussed in §5 together with solutions to handle them. Finally, some possible developments are presented in §6.

DOI: 10.13009/EUCASS2019-282

2. Analytical solution

A linear filter transfer function is an analytical function either of the Laplace variable p in the case of continuous time systems or of the $z = \exp(j2\pi f T_e)$ variable in the case of discrete time systems of sampling period T_e. As a consequence, its gain and phase are subjected to Bayard-Bode gain-phase relationship. Let choose a discrete time system in order to get periodic Bode diagram which can be completely described by one period (or even half a period thanks to symmetries). Let denote $H(f)$ the complex transfer function, G(f) its gain and ϕ(f) its phase, then

$$H(f) = G(f)\exp\big(-j\phi(f)\big) \tag{2.1}$$

where G(f) and ϕ(f) are real valued functions.
Taking the natural logarithm of the expression leads to

$$\ln H(f) = \ln G(f) - j\phi(f) \tag{2.2}$$

which is still an analytical function.

2.1 Hilbert transform

Basically, the Hilbert transform [2] can be seen as an operator able to create a phase shift of $-\pi/2$ applied to an input real valued function. For instance, starting from the function $\cos(t)$, the result of Hilbert transform is the function $\sin(t)$ and the result $\cos(t) + j\sin(t) = \exp(jt)$ is a strong analytical signal.
This means that starting from a gain function $G(f)$, the Hilbert transform of $\ln G(f)$ is $\phi(f)$. In other words, starting from a gain function, it is possible to generate a phase which is defined in a unique way under phase minimization filter assumption.
To illustrate, let consider a transfer function whose gain G(f) is such that

$$\ln G(f) = \lambda(\cos(2\pi f T_e) - 1) \tag{2.3}$$

It is a low pass filter whose static gain equals 1 and whose gain at half the sampling frequency equals $\exp(-2\lambda)$ (corresponding to $-8.686 \times 2\lambda$ dB).
Applying Hilbert transform, we get

$$\phi(f) = -\lambda\sin(2\pi f T_e) \tag{2.4}$$

As shown on Figure 1, the phase slope at low frequencies is proportional to the depth of the filter.

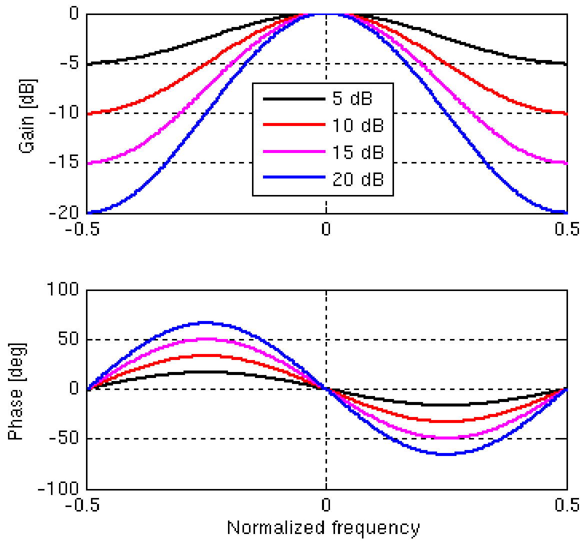

Figure 1: Cosine example Bode diagram

DOI: 10.13009/EUCASS2019-282

The Bode diagram and the Nichols plot can be seen as 3 projections of the 3D curve (frequency, gain, phase). To develop insight, it is useful to get used to the behaviour of the curve in 3D as presented on Figure 2 for the 20 dB filter. Here the gain is depicted by the green curve. The blue curve is the phase obtained by Hilbert transform. They can be seen as projections of the 3D black curve which represents the complex transfer function. The last projection (red curve) is the Nichols plot which is here a circle.
Conversely multiplying (2.2) by j leads to

$$j \ln H(f) = \phi(f) + j \ln G(f) \tag{2.5}$$

which is an analytical function as well.
Starting from a phase function $\phi(f)$ expressed in radians, one can then derive using Hilbert transform the natural logarithm of a gain function $G(f)$, which is defined in a unique way to an additive constant. That additive constant on the logarithm of the gain is equivalent to a multiplicative constant on the gain (multiplying a transfer function by a constant does not affect its phase).

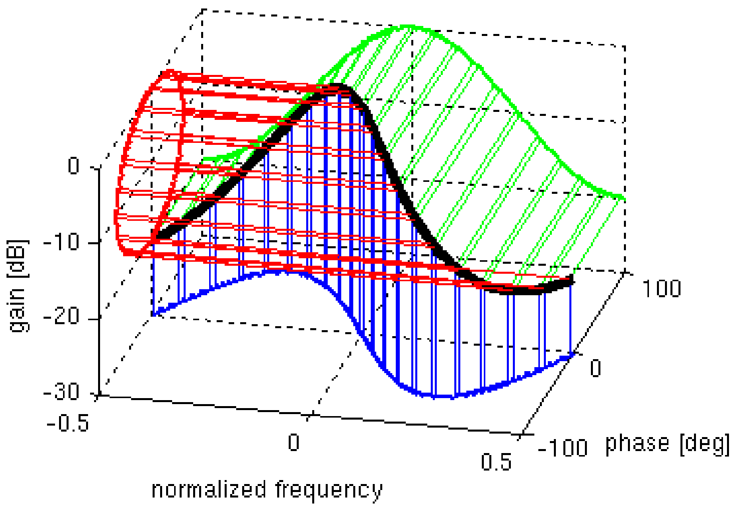

Figure 2: Cosine example 3D view

2.2 Iterative scheme

Gain and phase gauges: the filter design consists in finding an analytical transfer function that satisfy gain and phase gauges. For each frequency, the gain is constrained between a minimum gain curve (at least in the filter bandwidth) and a maximum gain curve (corresponding to the gain attenuation objective). Moreover, the phase of the filter might be constrained at low frequencies to minimize the equivalent delay of the filter or in some frequency ranges inside which a phase control is required.
According to Figure 3, the proposed Hilbert transform iterative scheme is based on the following process: starting from a gain curve equals to the maximum gain constraint curve, a phase is obtained through Hilbert transform. That phase curve is compared to the phase gauge and in the frequency range inside which the phase constraint is not satisfied, a new phase curve is built by replacing the obtained phase curve by the phase constraint. By construction, the new phase curve satisfies the phase constraints and a new gain curve is derived through Hilbert transform. The additive constant is adapted to normalize the gain (for instance the static gain for a low pass filter). Then the gain curve is compared to the gain gauge and again it is replaced by the gain constraint where it is not satisfied, and so on… At the end, the process converges towards gain and phase curves satisfying or not both gauges depending if the problem is too much constrained which means that there is no analytical solution at all (i.e. that Bayard Bode condition cannot be satisfied).

DOI: 10.13009/EUCASS2019-282

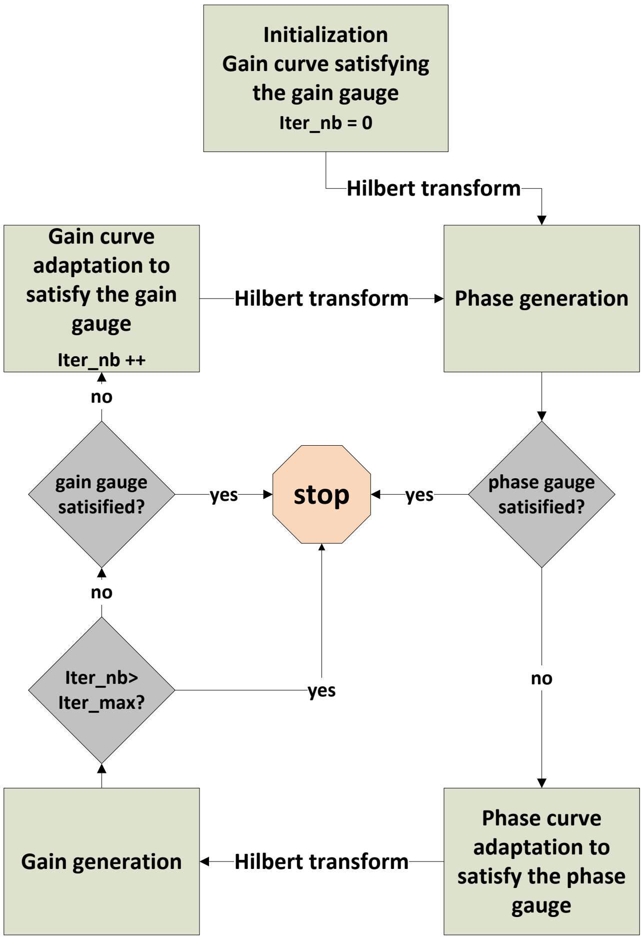

Figure 3: Iterative scheme organigram

3. Three illustrative examples

The approach can be tailored to a wide variety of filtering objectives, notably for the control or attenuation of poorly damped dynamics as frequently appear in large structures control loop such as launchers. It is then a complementary tool to the design of rigid body control laws [3].

3.1 1st example: Gain control with delay minimization

For this 1st example, the objective is to design a filter satisfying a gain gauge and minimizing its equivalent delay. The optimal solution could be directly obtained by initializing the iterative scheme with the maximum allowed gain and performing a single Hilbert transform. However, for understanding purpose, it is better to see how the iterative scheme is performing and this can be achieved using another initializing gain curve. The delay criterion is transformed into a linearly frequency dependence constraint at low frequencies phase. The maximum gain and maximum phase gauges are presented in Figure 4. If the phase frequency slope is high enough, the process converges (it is the case for the upper graph with delay equivalent to -12 deg at 1 Hz). We can even remark that there is still room for improvement as the final curve in red is lower than the gain gauge around gauge maximum. On the middle graph, the delay is equivalent to -6 deg at 1 Hz and for such a value, the iterative scheme does not converge (as can be seen around 1.5 Hz).
Then, to minimize the filter equivalent delay, a dichotomy w.r.t. phase frequency slope is performed. Finally, we find again (lower graph) that the optimal solution is the one for which the gain curve equals the maximum gain gauge.

DOI: 10.13009/EUCASS2019-282

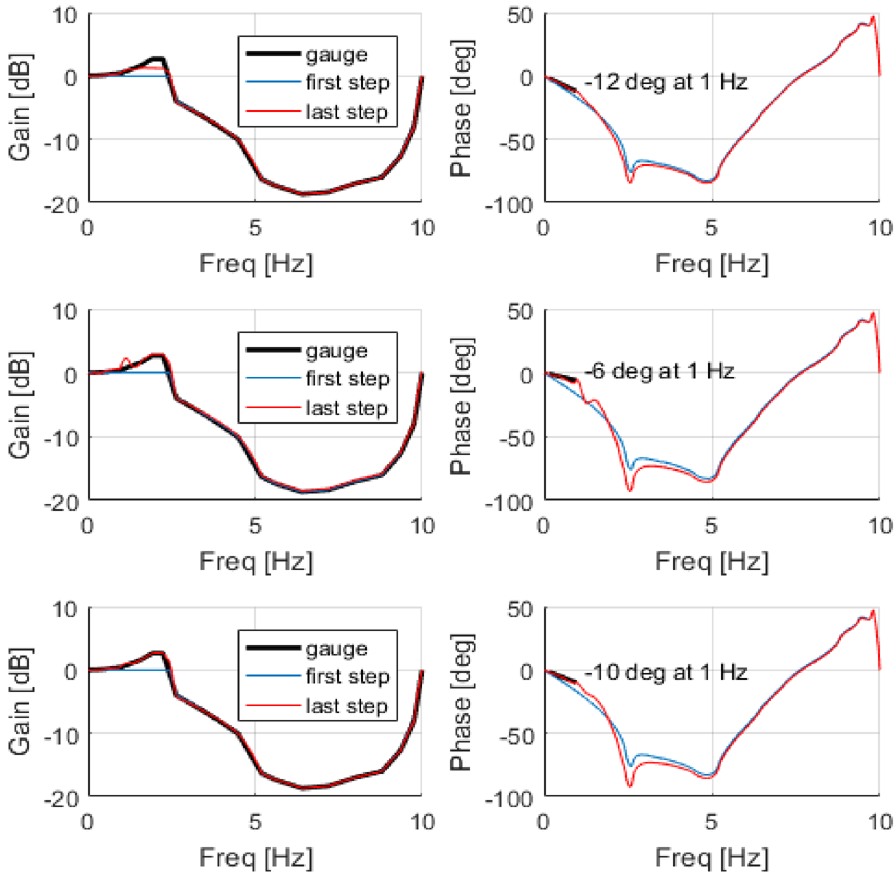

Figure 4: Gain attenuation with equivalent delay minimization

3.2 2nd example: Gain and phase control filter design with delay minimization

For that 2nd example, the objective is to perform a phase control for a low frequency dynamic (with a phase corridor to satisfy requiring a phase lead action) and to perform gain attenuation for higher frequencies dynamics. Here again the objective is to find a solution minimizing the equivalent delay and this can again be achieved through a dichotomy applied to the phase frequency slope imposed for the phase gauge at low frequencies.

After Wheeler's First Moral Principle, let discuss a priori the expected behaviour of the transfer function. The phase lead need will induce a gain increase in the neighbourhood of the phase lead frequency range. To minimize the equivalent delay, the optimal solution is to maximize the gain for frequencies higher than the phase lead action and this may also require a gain drop before the phase lead. Moreover, the gain drop will likely be located just ahead of the phase lead action whereas for lower frequencies, the gain will preferably be maximized.

We can check those predictions on Figure 5. The upper part corresponds to the gain and phase gauges. In some frequency range [3 Hz; 4 Hz], the gain is no more constrained but the phase is (the application requiring a phase control of a weakly damped dynamics). The initial step in blue leads to a phase curve which does not satisfy the phase gauge between 3 Hz and 4 Hz. Then the iterative scheme ends with the red curves for which both gauges are satisfied. In the lower part, the phase slope at 1 Hz has been minimized.

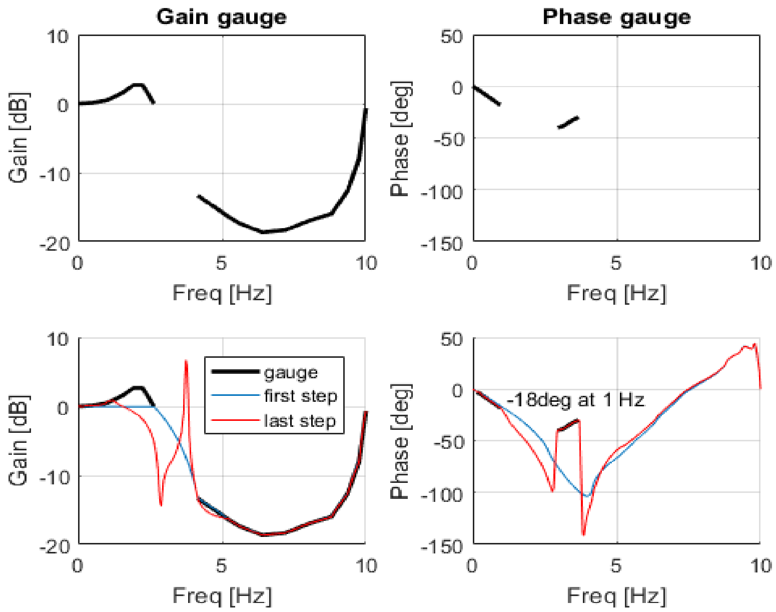

Figure 5: Gain attenuation and phase control with equivalent delay minimization

3.3 3rd example: Mixed Gain / Phase control with delay minimization

The 3rd example refers to a mixed gain/phase control of the same dynamic depending on the dynamic parameters scattering. Sometimes pure phase control or pure gain attenuation can become prohibitive whereas a phase control would not cost much for some parameters scattering as well as gain attenuation for other parameters scattering. An advantage of the proposed scheme is that logic operators can easily be taken into account during the curve adaptation steps. Precisely, it can be automatically decided to adapt the curve only if the gain gauge and the phase gauge are not conjointly satisfied.

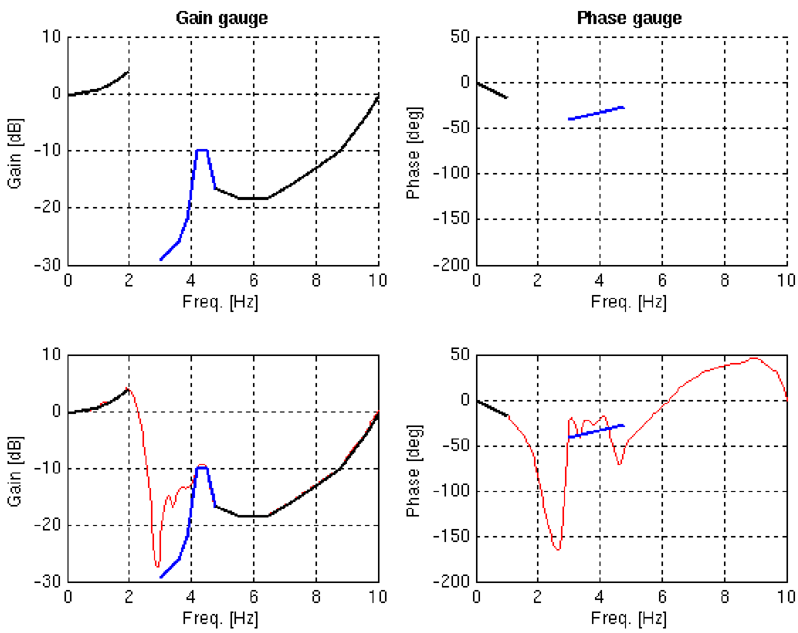

Figure 6: Mixed gain / phase control with equivalent delay minimization

Figure 6 presents the results with such a situation. Here phase and gain gauges (in blue) both exist in the range [3 Hz; 4.7 Hz] and any of the two gauges could be satisfied in that frequency range. The red curve presents the results at the end of the iterative process. Inside [3 Hz; 4.7 Hz] range, the gain gauge is satisfied up to 4.2 Hz, and above the phase gauge is satisfied. That transition frequency is not specified and arises from the process. Like for

DOI: 10.13009/EUCASS2019-282

previous examples, phase shift induced by the filter at low frequencies can be minimized by searching the minimum phase slope thanks to a dichotomy. Let note that the process is very quick (less than 1 s including the dichotomy loop).

4. Approximation by fixed order transfer functions

Analyticity is the only constraint which has been considered so far. Notwithstanding, to be able to implement the filter, one needs to find an approximation by a transfer function of fixed order. To perform this, a classical RMS (Root Mean Squares) method with weighted function is used. Figure 7 shows typical approximation results for different orders of the filter. In that case, an 8th order filter is already sufficient to avoid large deviations.

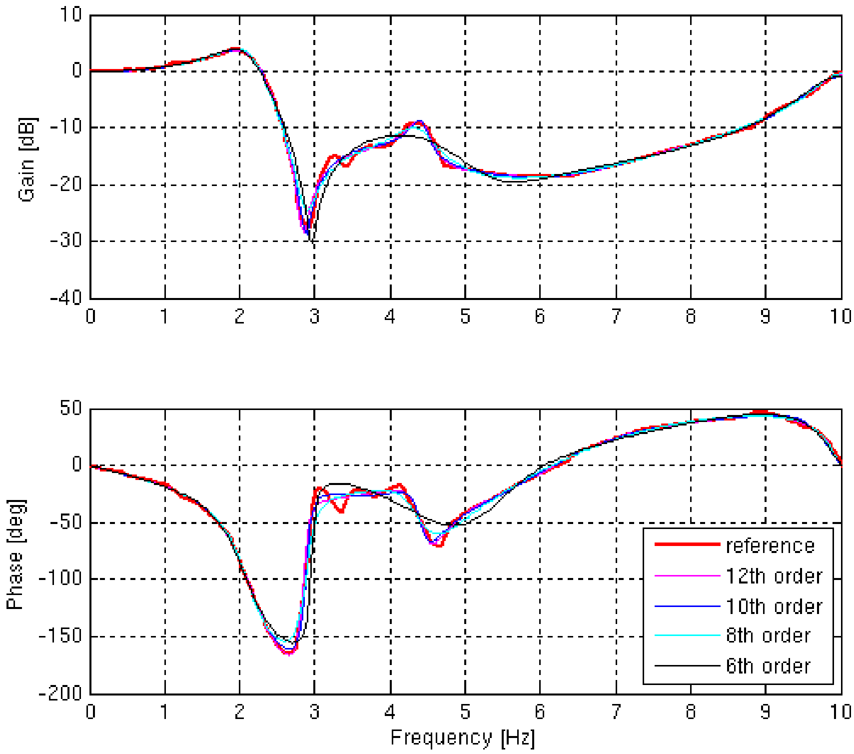

Figure 7: Approximation of an analytical solution by fixed order filters

5. Some drawbacks and how to tackle them

The approximation step may induce some local deviations. This can be corrected either by adding bias to the gauges or by implementing a descent algorithm using the result of the iterative scheme as initial condition to ensure convergence towards the local minimum.

Let note that unlike other classical fixed order methods, there is no a priori control of the structure of the filter. As a consequence, if the system to filter changes with time, the minimum order can even vary from one tuning time to the next one. This issue has to be solved. Preferred solution is here first to find minimum order filters, and then to compute its zeros and poles to perform a tracking of poles and zeros. To accommodate filter order transition, it is possible to add extra degrees of freedom by adding zeros and poles which compensate. It is then quite easy to perform a good tracking of poles and zeros and to get smooth evolutions that fit functional objectives. For instance, let consider two situations: in the first one, the plant presents a single dynamics which has to be attenuated using a notch filter whose notch frequency varies from 4 Hz to 6 Hz; in the second one, the plant present two dynamics at fixed frequencies 4 Hz and 6 Hz, at initial step the attenuation need is only required at 4 Hz and at final step it is only required at 6 Hz. The proposed process with extra degrees of freedom allows covering both situations. Figure 8 below shows a typical time evolution of a filter between two filters (called initial and final filters). On the left hand side, there is no extra degree of freedom but a simple interpolation of the single dynamics frequency (i.e. one biquadratic cell). On the right hand side, a second biquadratic cell is added and damping ratio are interpolated in way to compensate 6 Hz dynamics at initial time and 4 Hz dynamics at final time.

DOI: 10.13009/EUCASS2019-282

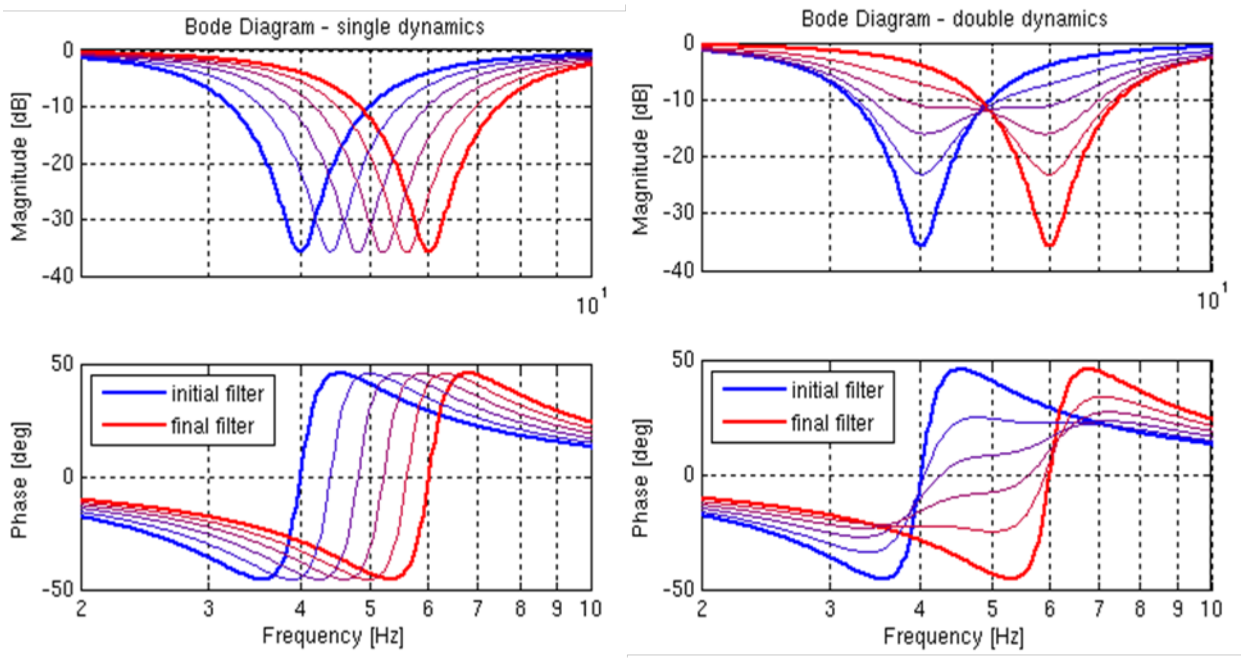

Figure 8: Filter interpolation capabilities

Figure 9 presents typical interpolated results when performing that kind of interpolation over time. It shows that even if a phase control is used or not depending on time value, it is possible to get smooth results between filters obtained using the iterative scheme and the approximation steps.

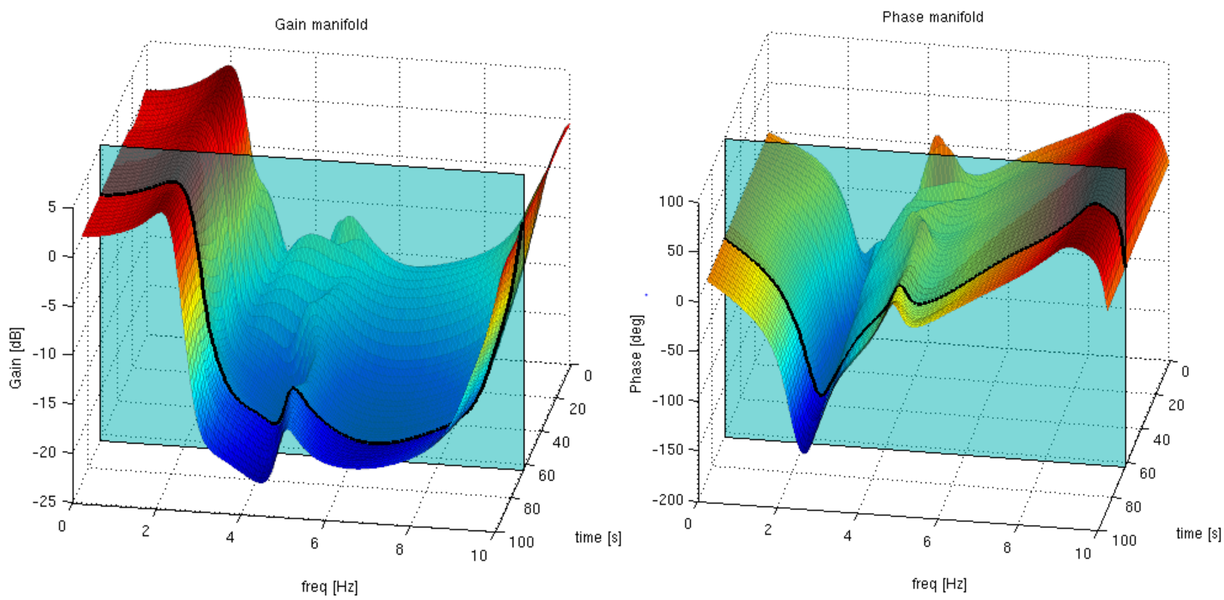

Figure 9: Example of filter interpolation over time

6. Further developments

There are different ways to further develop the proposed methodology, by applying it to other control objectives. Loop shaping is quite natural. The main evolution concerns the way constraints are expressed: gain and phase gauges (with potential logical conditions as presented previously) are replaced by a zone in the Nichols chart for each frequency. The objective is then to find an analytical transfer staying inside a 3D tube of constraints. Figure 10 presents a view of such constraints for a given frequency (red curve) compared to pure gain and phase constraints (dashed black lines). The same iterative scheme can be applied working alternatively on gain and on phase. At each

DOI: 10.13009/EUCASS2019-282

step, the new need gain (resp. phase) gauge is defined first by identifying frequencies for which the current transfer is not inside the tube and then by projection along gain (resp. phase) axis to reach the tube frontier. Main improvement comes from the ability to better trade gain and phase constraints (allowed domain is extended from green area to green and yellow areas).

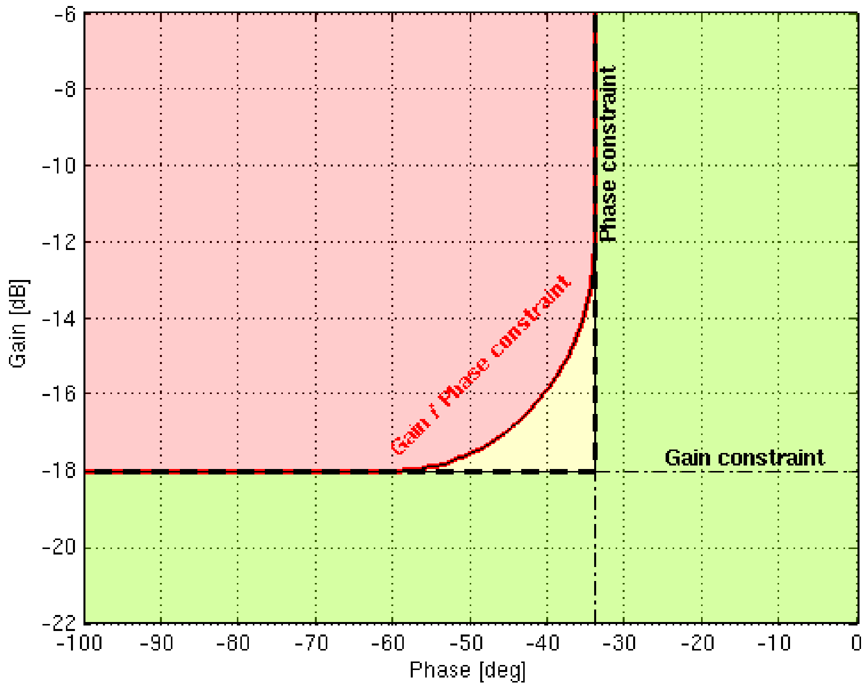

Figure 10: Gain and phase constraints at a given frequency

Here we can mention that the proposed approach can be combined with QFT (Quantitative Feedback Theory) developed by Issac Horowitz [4], whose objective is precisely to define the constraint tube taking into account plant structured uncertainties [5] and without any constraint about filter orders to describe the constraints. This means that there is no extra conservatism when using such an approach as could appear using mu-synthesis for instance.

7. Concluding remarks

This paper presents an iterative Hilbert transform scheme method helping control engineers to design filters and developing their insight of what should be optimal solutions. This approach, based on geometry to ease understanding, is quite different from tools using algebra to find optimal solutions. Moreover, this approach can be easily tailored and combined with QFT to cover a wide variety of filtering objectives.

References

[1] Taylor E. F., J. A. Wheeler, 1992, Spacetime Physics: Introduction to Special Relativity'. New York: W. H. Freeman.
[2] Hahn S. L., 1996, Hilbert transforms in signal processing, Artech House, Inc.,Boston.
[3] Saunois Ph. 2009, Comparative analysis of architectures for the control loop of launch vehicles during atmospheric flight, Aerospace Science and Technology 13, 150-156.
[4] Horowitz, I. M., 1973, Optimum Loop Transfer Function in Single-Loop Minimum Phase Feedback. Systems. Int. J. of Control 22, 97-113.
[5] Arnold, P. B., I. M. Horowitz, and C. H. Houpis, 1985, YF-16CCV Flight Control System Reconfiguration. Design Using Quantitative Feedback Theory Proceedings of the National Aerospace and Electronics Conference (NAECON) 1. 578-585.